UNITED
NATIONS

E

Economic and Social Council

Distr.
LIMITED

E/CONF.83/L.36
24 January 1991

ENGLISH ONLY

TWELFTH UNITED NATIONS REGIONAL
CARTOGRAPHIC CONFERENCE FOR
ASIA AND THE PACIFIC
Bangkok, 20-28 February 1991
Item 7 (b) of the provisional agenda*

CARTOGRAPHIC DATA RETRIEVAL, ANALYSIS, DEPICTION, PRESENTATION AND PRODUCT GENERATION: DIGITAL DATA BASES

Digital map data: A component of geographic and land information systems

Paper submitted by the United Kingdom of Great Britain and Northern Ireland**

* E/CONF.83/1.

** Prepared by M. Sowton, Head of Research and Development, Ordnance Survey, Southampton, United Kingdom and J. Price, Manager, Overseas Surveys Directorate, Ordnance Survey, Southampton, United Kingdom.

A map is essentially a compromise between scale, content and depiction, representing features in the terrain which have been surveyed, fair drawn and printed. Topographic maps are made for a variety of purposes and both content and scale are altered to fit the requirements of users. A map is of itself an information system containing a multitude of data which is transformed in the users brain into information.

With the advent of digital mapping techniques in the late 1960s it was anticipated that map production could be speeded up. What was not foreseen was the possibility that the data created for map production would be of value for other purposes.

By the 1980s the use of digital map data as a background for the display of map related information was rapidly increasing but at that time it was seldom required to support the analysis process.

The essential nature of a GIS is that it should be capable of linking data from various sources through spatial links contained in the data. Provided the spatial data also carries within it sufficient "intelligence" a computer system can use it to carry out analysis and make logical deductions. Improvements in the map data content has made this possible. LIS is defined in the paper as a subset of GIS.

Some examples of the way in which digital map data can be used in GIS/LIS applications are given.

Introduction

A conventional map is a sophisticated information system which has been measured by the surveyor, drawn by the cartographer and printed by the printer. Each works to a specification which determines the accuracy, the content and the area covered.

The specification determines which features in the terrain are selected and how they are measured, fair drawn and printed. Early cartographers embellished their work with pictorial symbols which represented significant features but whose main purpose was to overcome limitations of accuracy and content. By using designs to

fill the blank spaces surrounding the mapped area they produced an attractive result which was cartographically pleasing if metrically doubtful.

The scale of these early maps was determined to a great extent by the amount of detail to be shown and the positional accuracy of the survey had a lower significance than in modern cartography. Later, improvements in surveying techniques meant that more topographic detail could be measured and depicted accurately and the map scale took on more significance in the map design. As time passed the cartographer spent less time ornamenting his map and placed more emphasis on the correct depiction and placement of the topographic features.

As the ability of the surveyor to produce accurate positional surveys improved, the need for accurate large scale maps could be met. The cartographer was now faced with a dilemma; there was too much detail and the accuracy was greater than necessary for some of the map scales. Thus the cartographer had to develop rules which he applied to reduce the content and downgrade the accurate coordinates provided by the surveyor in order to produce smaller scale maps for specific purposes from larger scale surveys. This process, known as generalisation is a compromise between the map scale, the content and the way in which detail is depicted. For any particular map or more frequently map series, this compromise is arrived at by agreement between the map designers and the map users.

There is one further limitation imposed on the cartographic design and this is the sheet size. The sheet size is limited by the physical size which can be handled comfortably by the user, by storage considerations and to a limited extent by the size of the printing press. A number of standard printing sizes exist from which the map designer can select one which is suitable taking account of the effect upon ground coverage at the scale selected, sheet layout and design of the map series.

A map is an information system

The information content of any map depends on the purpose for which the map has been created. Thus a map for motorists needs to be at a scale where the limitations of printing size will allow sufficient area to be shown without being unmanageable, and the content will include significant features for navigation and a selection of the road network. If this is not sufficiently detailed a larger scale map has to be used but the penalty of less coverage and hence more map sheets has to be accepted. At the other end of the scale range are maps used for cadastral and planning purposes where scales sufficiently large to show street furniture such as lamp posts and manhole covers in their true position are used.

A map then is a representation of the real world where some selection of the content has taken place to service the needs of particular users which results in a specification which is a compromise between scale, content, accuracy, sheet size and finally the way in which the map is drawn and symbolised.

In order to use a map to unlock the information contained within it some skill in map reading is needed but once acquired it is a

relatively simple process for questions such as

- how far?
- where is?
- what is the distance between?
- how high is the ground at?

to be answered, while questions like

- can A be seen from B?
- what is the slope?
- which direction is the stream flowing?

can also be answered with more practice and some additional knowledge about maps. What cannot be answered without access to additional information are questions about the features in the map or related to them such as

- who lives at?
- how high is the bridge?
- what is the speed limit?

which are all simple examples of the kind of information contained in a Geographic Information System (GIS).

Thematic Maps

Between the topographic map and the GIS is a whole range of purpose built thematic maps which have been designed to portray information related to geographic positions and sometimes printed with the topographic map as a background. Examples of these are geological and soil maps. Demographic maps are often designed with symbols representing the quantitative information related to a geographic position, such as population to urban centres. This may be done using increasing symbol sizes. However there is a drawback to this display as any breakdown of the information requires yet another map, so that for information about the distribution of ethnic groups in the urban centres or the relative numbers by age or sex this becomes a cumbersome procedure. It is for this purpose that the GIS is so effective.

Digital Mapping

Before looking at GIS it is necessary to consider how a computer differs from a person when considering a map. People have a perception of maps which is affected by their knowledge of the real world, their experience and their intuition which interacts with the information within the map giving its relationship to the real world such as scale, orientation, symbolisation, projection and degree of generalisation. The two latter items affect the perception by introducing degrees of distortion and displacement which complicate

the way in which the map is interpreted. Finally people have to contend with their preconceived ideas and even prejudices which distort the perception. Nonetheless the human brain is extremely efficient at extracting information from the huge amount of detail available in a topographic map.

Questions about relative position, slope, size, distance, and connectivity are rapidly answered whereas a computer would need to be given details which a person can rapidly infer, such as that there is no logical connection between a stream and a road even at a ford or that a motorway only connects to other roads at specific junctions.

However a computer programmed correctly and given the correct data can rapidly overcome differences which arise from projections and to some extent generalisation, which many people fail to understand.

In order to do this the data must include information which the computer program uses to distinguish the various parts of the map.

- geometry
- reference systems
- feature codes
- attributes
- topology.

Before looking at the way in which these are used it is worth considering the way digital mapping at the Ordnance Survey (OS), the national mapping agency of Great Britain, developed.

In the late 1960s it was thought that if digital processes could be effectively introduced to produce maps automatically from stored data it would be possible not only to emulate the existing map specifications but also through software to produce generalised maps at a variety of scales from the same data. During these early investigations there were few demands for map data beyond the map making processes so development concentrated almost entirely on using the data for map production.

Although some work was done on small scales, investigations concentrated on the production of digital data for the large scale urban mapping and the associated generalisation problems.

By the early 1980s a demand for digital map data was emerging for use as a background reference map for the display of information predominantly from the digitised records of cables or pipe lines of the public utilities. Data for this purpose did not require all the various coding which had been introduced for cartography and as a result, some reduction in content was introduced in order to make the data more acceptable. From about 1988 a demand for more sophisticated data has emerged partly to improve cartographic processes and more specifically to service GIS. However the predominant use of digital data in Great Britain remains at large scale (1:1250) for the recording of utility distribution networks, and at local government level where there is an increasing number of municipal authorities developing GIS/LIS for urban land management and physical planning. Although, investigations into the wide range

of possible applications in GIS are beginning to create demands for data at other scales.

Geometry

At the lowest level a digital map is composed of the coordinates which define positions on the earth's surface for all the features included in the data set. In its simplest form the coordinates which represent these features are collected in any order and the resulting strings of coordinates provide the information which the computer needs to reproduce a map. At this level there is no intelligence in the data, thus one coordinate string cannot be identified from another nor are the points where strings or lines cross specifically defined, unless the crossing point has been coordinated as a deliberate act.

Reference Systems

In order to introduce a relationship between coordinated points on the earth's surface and their representation as a map, computers rely on various reference systems though which scale and projection are related to the real world in exactly the same way that a conventional map establishes these relationships. The difference is that the conventional map series has the scale and projection determined as part of the specification whereas computers are capable of varying projection and scale rapidly and accurately. In practice a coordinate reference system is selected for a particular country and application which here in Great Britain is the National Grid but a system based on latitude and longitude or Universal Transverse Mercator could be used with equal effect. Provided there is a relationship which can be programmed into the computer system transformation of coordinates between reference systems can be carried out. For a computer the concept of scale is not significant since scale is dependent purely on the relationship between the screen display and the real world. What is of much greater importance is the resolution of the coordinates which have been captured since coordinates captured to the nearest 5 metres may be quite adequate for a scale of 1:50 000 whereas at 1:1000 they do not have sufficient accuracy for that resolution.

In addition to the coordinate reference system there are other ways of selecting locations in the digital data such as addresses and postcodes but these are dependent on being included in the data set with suitable pointers to the coordinate system so that they can be recognised by the computer program.

Feature Codes

To unlock the relationship between a line in the data and a feature in the terrain the computer needs to be told what the coordinate strings represent and this is the basis through which the digital data takes on a more coherent form. Each line in the data is assigned a feature code which identifies what it represents on the ground. Thus a building outline can be separated from the fences which bound the land parcel and the edges of a road can be identified as being made up of a series of lines, coded building, fence, ditch etc through linking them to the centre line which acts as a pointer to the edges. In a similar way point objects can be

coded so that the correct identification can be assigned and the correct symbol displayed.

Attributes

Having located the lines through coordinates and identified what they represent, additional information can be assigned through seed points in the data. In a cartographic context these attributes would be the line styles such as pecked lines for road curbs or line weights depending on the specification and would include road names and house numbers. Attributes can arise from a number of sources and, in a GIS, seed points in the map data are linked to non-spatial data in associated databases to give the geographic relationship to that data.

Topology

In order to perform properly, within a sophisticated information system the data needs to be structured in such a way that the computer can readily determine the logic in the data, while still relying on the geometry and the feature codes to define the features. In topologically structured data the component parts of the data are identified so that significant junctions are made into unique points called nodes while the lines between the nodes regardless of whether sinuous or straight are termed links. Only where two links cross which have a significant connection in the real world are the links broken and a node created. Thus where a power line crosses a fence the two links would cross without a node at the crossing point. In topology the data is formed into polygons and seed points are created so that attributes can be related to the objects formed by the polygons. This concept allows the inside and outside of a polygon to be separately identified and linked to separate attributes. It is now possible to identify objects in the data so that a collection of building polygons with a perimeter and possibly separate but associated polygons, can be identified as a school. It is a short step from here to catchment analysis given the school pattern and the location of the homes of school age children or even the planning of school bus routes to transport the children to school.

In addition to the polygons formed from the links, topology makes possible the identification of networks so that either river or road networks can be formed for analysis in relation to the attributes which can be attached to the links.

Finally there are relationships which identify the connectivity so that the link which passes under can be identified in its spatial position relative to the link passing over.

A prime example of a topological map is probably familiar to most people who visit London, without realising the significance of the diagram. The London Underground map has no geometric framework, it has variable scale and the positions of the stations only very approximately give their relationship to other stations particularly those on another line. However what is shown is the network and the connectivity which allows a route to be planned but which, without some knowledge of the true shape of the layout in relation to the city streets, may not always result in the quickest journey. A GIS

application could very simply be designed to produce such a result if journey times were attached to each link and coordinates were assigned to the stations to relate them to a suitable topographic map.

Levels of use of digital data

There are three main levels of use for digital map data. In the early days of digital mapping data increased in complexity from level 1 to level 3, but recent improvements in map production can only be achieved with data which satisfies many of the requirements of level 3.

Level 1 - Data for the production of maps and plans.

The use of digital map data makes it possible to produce maps to a variety of scales and content. The simplest form of data which includes geometry and feature coding is capable of map production to a specification which does not include symbols or ornamentation. However fully automated production can be achieved through the improvement of the data by using polygons which contain seed points to indicate where symbols and ornamentation are to be included. Thus woodland, slopes and building stipples can now be plotted in their correct style and position automatically. This more versatile data makes it possible to eliminate the constraints of sheet size, map layout and within limits, scale. It also allows the map design to be changed through alteration of the symbols and the addition of colour, and with special software can carry out a certain degree of generalisation automatically.

Level 2 - Background to user data.

Even using the simplest form of digital map data will allow information such as pipeline routes to be recorded against the digital map background. The result can be displayed on a screen or plotted as a composite document, but until additional intelligence is included in the map data it is only possible to analyse the associated user data to the level of intelligence in that data. Thus it would be possible to obtain information about the pipeline network such as pipe sizes, material flow, etc, but any underlying information relating to the customers being supplied could not be accessed through the map data without a more sophisticated map dataset with associated attributes which would be essential for GIS applications.

Level 3 - Component of a GIS/LIS.

Although level 2 is a simple GIS, in order to be used as a component of a GIS/LIS the map data would need to include sufficient topology to make the attachment of both graphic and non-graphic attributes possible. In such an application it would be possible to relate the digital map data to associated databases which would allow various separate sources of information to be related through the spatial reference system of the map data. In the case of utility records it would be possible to identify the house number or name, the occupier, consumption, billing details, etc.

Geographic Information Systems (GIS)

There are numerous definitions of GIS but all err on the side of brevity leaving out as a result many of the functions of GIS by not specifying exactly what it includes.

The essential nature of a GIS is that it should be capable of linking a variety of data through a common spatial reference system. A GIS allows the collection, updating and display of a number of unconnected data sets bringing them into a common reference system for spatial analysis from which relationships can be identified and decisions made. Through the structure in the data it is possible to link non-spatial data to the polygon seeds and to individual links in the map data and to manipulate the various data sets as one data set.

A GIS is not concerned primarily with maps. While a map may be a way in which to access the data and simplify storage it is the spatial reference system which allows the analysis and integration of datasets to happen effectively. However when it comes to the display of results there is no better way in which to show location and relationship between sites or events than the digital map. When this is done it is important to produce a clear and concise map to a high quality which in turn adds credibility to the results. The structure and feature coding in the map data allow the data to be manipulated to produce maps which are suitable for display.

Finally the map has various components which when attributes have been attached, can be used to provide analyses of polygons or networks. Special subsets of the map data can be used to create specialised data products to support specific GIS applications. When used in these ways the map data takes on a high level of importance as a GIS component.

Land Information Systems (LIS)

For the purpose of this paper an LIS is defined as a subset of GIS attributes. The principal difference being that an LIS uses large scale digital map data as a base on which to overlay land related digital information in a computer system for the specific purpose of managing land related activities. This usually requires digital map data captured at large scales (1:1000 - 1:10 000) in order to provide a topological base to show the relative position and relationship of buildings/land plots. The definitions of the previous paragraph in bold type apply also to LIS.

OS Achievements with Digital Map Data

By 31 March 1991, the OS will have completed the digitising of all published national plans at 1:1250 scale, (this scale of mapping covers urban centres with a population of more than 20,000). The digitising of 1:2500 scale mapping covering the smaller towns and populated rural areas is also proceeding rapidly, and is 20% complete.

In order to make this structured digital data available to all potential users for their GIS/LIS use, the OS has developed the National Transfer Format (NTF) which is capable of transferring data

As an example of GIS in a planning role the digital map of an area shows the field boundaries, road layout, buildings and some information about the land use such as woodland, orchards, marsh etc. If the pattern of land ownership is added, together with information about the existence or otherwise of planning permission it is possible to determine the development potential and suitability of various sites on offer. The map background makes it possible to display the various analyses and also can provide details about access and parcel areas which together with land values would provide a good basis for making a decision about a development proposal. However simple it may appear the components of a GIS have to be collected and integrated into the GIS, thus in this example unless the ownership of the land can be established and which parcels already have planning permission the initial stage in the process must be data collection and this should not be underestimated in either time required or cost to do it.

The previous example has illustrated large scale map data applications which show an overlap of GIS and LIS, but a GIS is equally suited to the consideration of problems over wider areas where factors affecting the environment can be studied in relation to their effect on the population. One such study by a Regional Research Laboratory is concerned with the investigation of the reasons why clusters of cases of childhood leukemia occur. For this study the map base needs to be of sufficient resolution to show the location of population centres significant to the study over the whole area and yet be suitable for display of smaller areas at larger scale.

The population census data being collected in 1991 for the UK will create the requirement for map displays based on the small scale maps of the country combined with the ability to display data related to local government wards, census enumeration districts and possibly unit post codes in the urban areas. Thus the generalisation inherent in the maps will have to be introduced into the non-graphic data to allow macro spatial analysis to be carried out from the detailed data which the census will provide.

In a similar way data for in-car navigation is a combination of the macro route network and the very detailed road systems in the urban areas. Attributes relating to the road network which include direction restrictions, speed limits, bridge heights, width restrictions etc which are of a permanent nature or change infrequently, and those of an ephemeral nature such as lane restrictions, road works, accidents etc are linked to relevant points in the map data to create a complete database of the road network which can be used for a variety of applications ranging from in-car navigation to route-planning. Clearly a database of this sort combined with traffic patterns and volumes would be able to answer questions such as - what happens if main interchanges on national highways are closed on national holidays?

LIS Applications in the UK

At the municipal level, Milton Keynes Development Corporation has created a corporate land information system which is used to keep all the land transactions and associated records of occupancy, ownership, leasehold etc.

With reference to Land Title Registration, both HM Land Registry in England and Wales and the Registers of Scotland now have active programmes to transfer title records from the existing OS cadastral plans onto OS digital map data.

In Great Britain, because the national digital map data is both topographic (geographic) and cadastral (land) there is a tendency for GIS and LIS to overlap.

Conclusion

The foregoing has given a very brief description of some of the considerations concerning digital mapping, its relationship to GIS/LIS and developments and applications in the UK. There are many other factors for which there is insufficient space and time to give a mention. It must be remembered that for a GIS/LIS to be effective it needs sufficient data for the investigation in hand and where data is not available its collection is a major consideration.

Not only is data a major resource but its management and maintenance are of prime importance. Out of date information will give flawed results but an out of date background map is of equal concern. Revision of the map base affects the whole GIS/LIS and it is absolutely vital that the way in which this is carried out is carefully considered as part of the GIS/LIS design.

Selected Bibliography

Principles of Geographic Information Systems for Land Resources Assessment, Professor P A Burroughs, Oxford University Press 1986.

Geographic Information Systems: Overview, Principles and Applications edited by Maguire, Goodchild and Rhind, shortly to be published by Longman Scientific and Technical UK.

Cartographic Design and Production 2nd Edition, John S Keates, Longman Scientific and Technical UK 1989.

Land Information Management, P F Dale and J D McLaughlin, Clarenden Press, Oxford 1988.